Uncertainty Analysis to Better Confront Model Uncertainty

Paul K. Davis, Steven W. Popper

Pardee RAND Graduate School

WR-1272
January 2019

RAND working papers are intended to share researchers' latest findings and to solicit informal peer review. They have been approved for circulation by Pardee RAND Graduate School but have not been formally edited or peer reviewed. Unless otherwise indicated, working papers can be quoted and cited without permission of the author, provided the source is clearly referred to as a working paper. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors. RAND® is a registered trademark.
Confronting Model Uncertainty

Paul K. Davis and Steven W. Popper
RAND Corporation and Pardee RAND Graduate School
1776 Main Street; Santa Monica, CA 90407-2138
December 21, 2018

Corresponding author: Davis (pdavis@rand.org; 310 451 6912; ORCID 0000-0002-2913-2168)
Abstract

It is increasingly recognized that good policy analysis requires addressing uncertainties seriously. We argue that such analysis needs to be much better than in the past in routinely addressing not only uncertainty about model inputs, on which much progress has been made, but also uncertainty about the model itself and major disagreements of perspective. Doing so is feasible and studies should be organized accordingly. Success will depend on analysts understanding how to proceed and policymakers demanding that they do so. This will require major changes of culture that will likely occur only over time. In the future, analysis that fails to address effectively uncertainties about both models and their inputs will be regarded as fatally flawed.

Keywords

Uncertainty analysis; deep uncertainty; policy analysis; model uncertainty; structural uncertainty
Confronting Uncertainty

Paul K. Davis and Steven W. Popper

1 Introduction: Analysis When the Model is Uncertain

1.1 Objectives

This paper is written for the decision- and risk-analysis communities, policy analysts more generally, and for the policymakers who commission such studies and analysis. Although addressing uncertainty analysis more broadly, we highlight a specific challenge: “How do we conduct policy analysis when we lack confidence in selecting the correct model, i.e., when we cannot agree on or are unsure about how the world works?” This differs from parametric uncertainty, i.e., uncertainty about inputs to the models. Methods now exist and have been applied in numerous cases for dealing with parametric uncertainty. But what if the model itself is problematic? That constitutes model uncertainty (also called structural uncertainty). Good policy analysis needs to confront it far better than in the past. The issue is longstanding (Quade, 1968; Morgan & Henrion, 1992; Davis, 1994; Lempert, Popper, & Bankes, 2003) and good analysts often recognize the issue. Nonetheless, it is common ultimately to place bets on the one model that somehow seems best or most acceptable. We urge a different approach that explores model uncertainty as a primary facet of analysis. We suspect that analysis that fails to do so will in the future be seen as fatally flawed.

In stage-setting, we note that policymakers not uncommonly recognize that the quantitative and seemingly authoritative analysis with which they are presented rests on fundamental assumptions that are arguable at best. Their reticence to accept the implications derived from such analysis is then warranted—it is a sign of wise skepticism rather than parochialism or an anti-science perspective. Their skepticism might be ameliorated if analysis dealt transparently with fundamental uncertainties and disagreements. Current examples of where this might matter include evaluation of proposals to reform the tax system (will the recent tax cuts pay their way by stimulating sustained growth?) and evaluation of policies to mitigate the consequences of climate change (will suggested measures be effective enough to pay their way?). To be sure, some individuals will be impervious to analysis if they don’t like the implications, but we have in mind those who would engage, interact vigorously, and be influenced if only the analysis were more intelligible and helpful. As a side note, we suspect that better analytical practice will elevate the stature of analysis (and analysts) with policymakers.

1.2 Structure of Paper

In the remainder of this paper we discuss why it is important to address model uncertainties (Section 2), why doing so is more feasible than conventional wisdom would have it (Section 3), what a way ahead would look like analytically and procedurally (Section 4), how past projects can be seen as existence proofs justifying the way ahead (Section 5), and our
conclusions and recommendations (Section 6). Our examples draw on instances of economic and otherwise strategic choice under uncertainty. Some highlight issues related to the psychology of human reasoning and the complexities of human and social behavior.

2 Importance of the Problem

How important is it, really, to assess the kind of deep and mysterious model uncertainties discussed above? Three examples illustrate the importance.

2.1 The Financial Collapse of 2008

The financial crisis of 2008 was profound with aspects of structural collapse. Although the economy has recovered, America continues to suffer the consequences of a lost decade with economic production and job creation far below their potential. The origins of crisis were complex, but a major factor was misrepresentation of human behavior in policymakers’ mental models and the corresponding financial risk models used in business and government as discussed in official, scholarly, and popular accounts (Financial Crisis Inquiry Commission, 2011; Thaler & Sunstein, 2008; Krugman, 2009; Lewis, 2010). The previous head of the Federal Reserve, Alan Greenspan, famously testified in Congress that (U.S. House of Representatives, 2008):

“Those of us who have looked to the self-interest of lending institutions to protect shareholders’ equity, myself included, are in a state of shocked disbelief.”

Before the crisis, lending institutions took much greater risks than were prudent, behaviorally incentivized by the need to compete frantically with other lenders promising high rates of return, by the potential for extraordinary personal gain, and by the belief that housing prices would always rise (Financial Crisis Inquiry Commission, 2011, p. 6). Perhaps they assumed that the government would bail them out if need be, or that they could move in or out of the market before problems arose. In any case, the risk calculations were a poor match for reality. At root, the calculations incorrectly assumed rational-actor organizations and that market forces would correspondingly mitigate or eliminate irrationalities. The admonition about “Garbage in, garbage out” applies not only to the data feeding a computational model, but also to the conceptual model behind it.

2.2 The Limits To Growth

Another classic case provides a second example. Details of model-based analysis seldom receive public attention, but a tumult followed publication of The Limits of Growth in the early 1970s (Meadows, Meadows, Randers, & Behrens III, 1972). The authors were associates of famed MIT professor Jay Forrester, who introduced the method of System Dynamics. The book warned that modern civilization would collapse in the 21st century due to increasingly unresolvable environmental and economic issues if trends continued. Some criticisms were strenuous (Cole et al., 1973) with Forrester responding vigorously (Forrester et al., 1974).
Later exchanges were more restrained (Nordhaus, 1992), but the debate was imperfect (Bardi, 2011).

The computer model employed depended on equations embodying assertions about demographic, economic, technological and political processes. To be sure, parameters within those equations were treated as uncertain, but not so much the processes themselves. In particular, the model incorporated exponential growth of population (so-called Malthusian growth), although not in the naive way assumed by critics. Arguably, the model did not allow adequately for revolutionary changes (Kahn, 1983), as in technology, energy production, agriculture, and human behavior (e.g., birth rates). The authors did not anticipate the drastic changes that have occurred in the energy business, some due to dramatic innovation (e.g., solar energy, tertiary recovery) and some to effective use of previously lesser sources (natural gas). Arguably, their model did not allow market forces to respond adequately as resource scarcity grew. In addition, the model did not distinguish well among different regions and climates, which some critics saw as a major shortcoming.

The authors were sometimes seen as extreme advocates for environmentalism and slow growth, but the book had much to say that was important and provocative. The prevailing conventional wisdom in the 1980s and 1990s was that the Limits to Growth work had been discredited. In more recent years, however, several researchers have found remarkable agreement between empirical reality over time and many of the model's predictions in its business-as-usual scenario (Turner, 2008; Bardi, 2011). Other authors have noted social and intellectual analogies between the debate about Limits to Growth and more recent debates about climate change (Eastin, Grundman, & Prakash, 2010).

A lesson from the Limits to Growth saga is to view computer models as something to use in exploring and debating, rather than in confidently predicting. The book flashed valid danger signals suggesting how bad the future would be unless major changes occurred. The contingent nature of this discussion was more clearly expressed in the 30-year update (Meadows, Randers, & Meadows, 2004), but was present from the outset. As Meadows observed in a 2012 retrospective, however (Gamino, 2012), policymakers and people generally have a diffi-

---

1 Some model uncertainties could be studied by varying several model parameters jointly, as noted in the authors' response to criticism (Cole, Freeman, Jahoda, & Pavitt, 1973, p. 223; Forrester, Low, & Mass., 1974). Doing so was not intuitive to critics unaccustomed to system dynamic models. For example, the net population growth rate (something that might seem like an input assumption) was an output of the system dynamics model and was influenced by a number of factors. One factor appeared superficially to be “the” assumed growth rate. Instead, it was actually just one contributor, a highly conditional growth rate.

2 Meadows noted the primary criticisms in a forty year retrospective (Meadows, 2012). Another criticism was that the slow- or no-growth movement of the 1970s tended to assume that energy production would grow in proportion to economic growth. In recent times these have become decoupled for the United States (energy production has been flat, despite economic growth). Worldwide, energy production has increased with economic growth, but much less than proportionally (Business Council for Sustainable Energy, 2017).
cult time with contingent predictions. He went on to claim that no one ever seemed to have looked at the alternative scenarios. No, collapse was not inevitable, and the study said so.

At the time of the *Limits to Growth* book, scientists were doing well if they ran the simulation with varied assumptions about the parameters collected as alternative scenarios. To also vary the underlying conceptual models, and certainly to do so in a comprehensible way enabling productive debate, was for the most part a bridge too far. More is feasible today but doing it well is a grand challenge for policy analysts.

### 2.3 The 2003 Iraq War and the 2007 Surge Operation

As a last example of why models matter, consider a strategic-military decision with profound consequences. The primary decisions behind the invasion of Iraq in 2003 and its conduct did not depend on complex models and analysis. They did, however, depend on assessments of what would happen if the U.S. and allies of convenience mounted such an invasion. The decisions reflected selective insights from relatively simple war gaming and other tools of analysis.\(^3\) Such background work allowed President Bush to believe that the military operation could be fast and decisive, toppling the regime of Saddam Hussein and defeating its armed forces in short order. Experience in the first Gulf War of 1991 added confirmation: there was no way that Saddam's military could defeat American military forces on the battlefield. Such matters were also played out in computer simulations.

Putting aside the fact that some premises underlying the war proved wrong (Saddam did not have nuclear weapons), the actual experience demonstrated that the planners' conceptual model of conflict was also wrong (Kaplan, 2013; Gordon & Trainor, 2012). They had expected conflict would be short, decisive, and easy. What emerged was insurgency using guerrilla tactics and low-level civil war. The models and war games were accurate enough for planning initial military operations, but proved to have little to do with how matters played out later.\(^4\)

In 2007, after years of dismal results, President Bush authorized the Surge, a last-ditch U.S. effort to salvage what had become a losing war. Most diplomats, military leaders, and analysts opposed the surge, believing that it would either fail or be counterproductive (Gordon & Trainor, 2012), but President Bush was persuaded by arguments that stemmed from a group of analysts led by Frederick Kagan of the American Enterprise Institute and associates. They conducted map games that could foresee a road to victory. They also generated quantitative estimates of force requirements. The conceptual model assumed that decisive military force was only necessary in selected places in Iraq. Their concept convinced then Colonel

---

\(^3\) Historical analysis flashed warning signs about an invasion with such a small ground force (Quinlivan, 1995). Reflecting such experience, Army Chief of Staff Eric Shinseki told Congress that a much larger ground force was needed, something immediately derided by political leaders.

\(^4\) A partial exception was experience with the political simulation model, Senturion, which combines some deep political science (Bueno de Mesquita, 1983) with structured expert inputs. The analysis anticipated aspects of what transpired in Iraq after the 2003 invasion and after after the 2005 Iraqi election (Abdollahian, Barnick, Efird, & Kugler, 2006)
H.R. McMaster and retired Army General Jack Keane, who helped sell it to the President and Vice President (see Chapter 24 of Gordon & Trainor, 2012. A distinctly human factor in the decision was President Bush's conclusion that it was unacceptable to signal defeat without trying every last card in the deck.

As it happened, the subsequent operations were well conducted and successful under the leadership of General David Petraeus, although the surge in U.S. forces was only one of several reasons for the temporary success. The "Anbar Awakening" played a major role—the Sunni tribal leaders in Anbar concluded that it was in their interest to collaborate with Americans. Another factor was ethnic redistribution as residents of mixed neighborhoods voted with their feet during the quasi-civil war. Also, President Maliki reluctantly recognized that Shia militias were a threat to his own power and allowed actions that he had previously prohibited.

Detailed computer modeling performed in this period represented only force movements and military rules-of-thumb calculations—not the longer-run issues that were more quintessentially political, such as the parochial shortcomings of Nouri al-Maliki that led to squandering opportunities that had been created.

U.S. troops departed over the period 2007-2011 as promised in a formal agreement reached under President Bush, which President Obama honored despite calls for leaving forces in Iraq. Subsequently, Iraq fell again into sectarian struggle and the Islam State emerged to fill the vacuum. As of late 2016, the United States began increasing its force commitment again, although not dramatically.

This case, again, illustrated the profound importance of questioning the conceptual model behind a string of reasoning, not just the values of parameters in a formalized model.

3 Myths and Realities about Uncertainty Analysis

Section 2 illustrated why analysis should consider model uncertainties (something readers already know). For as long as the authors can remember, however, it has been widely believed that directly confronting model uncertainty comprehensively is not actually feasible. The most pragmatic approach, it is often believed, is to make reasonable assumptions and plunge on: i.e., it is better to try something and adapt later than to suffer paralysis by analysis. The stance once had considerable basis in practical experience, but it is now a myth that needs to be refuted: uncertainty analysis—even about model uncertainties—can be very helpful rather than paralyzing.

Uncertainties can now be dealt with by employing more sophisticated approaches and policymakers can be assisted in making informed, well hedged decisions, something empha-

---

5 The withdrawal of forces in 2011 remains controversial. The 2008 agreement called for that total withdrawal. Some believe that a significant force presence could have been maintained had the U.S. more energetically pressured the Iraqi government (Brennan, 2014). Others believe that Iraqi politics did not permit a change large enough to accomplish anything, a view reportedly supported by an assertion of Maliki in 2011 (Korb, 2015).
The classic work on uncertainty analysis remains excellent (Morgan & Henrion, 1992). We ourselves have been heavily involved in two streams of work on the subject since the 1980s. One has focused on defense planning (Davis, 1994; Davis, 2014). The other began with a rethinking of modeling for policy analysis (Bankes, 1993) and proceeded with work on climate change (Lempert et al., 1996). That second stream has expanded to address diverse issues with the methodology called robust decisionmaking (RDM) (Lempert et al., 2003; Popper et al., 2005; Lempert et al., 2006). Interest in decisionmaking under deep uncertainty now has strong international interest, such as efforts at Delft University in the Netherlands (Haasnoot et al., 2013) and as evidenced by a vibrant Society for Decision Making under Deep Uncertainty (DMDU; http://www.deepuncertainty.org). Exceptions exist, but most of the research in these two strands has dealt with parametric uncertainty.
### Table 1 Methods for Addressing Model Uncertainty

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
</tr>
</thead>
</table>
| Scenarios             | Use different scenarios to group alternative assumption sets, sometimes with the sets differing so much as to effectively change the model.  
  *Example:* strategic-planning that describes the world as potentially unfolding by a no-surprises scenario, a scenario with surprising technological breakthroughs and economic growth, or a scenario with relative stagnation |
| Competitive Models    | Use competing model alternatives. Often, the most important disagreements or other uncertainties can be grouped into two or three clumps. Many variations may exist within each, but the differences between clumps is more important. If so, one can build representative versions of each clump and compare results. This sharpens issues.  
  *Example:* in contemplating coercive strategy, evaluate with alternative cognitive models of how the adversary reasons. Is the adversary fearful and merely attempting to deter or is the adversary contemptuous of others and planning aggression against the weak? [a] |
| Bounding Models       | Identify models that bound the range of not-implausible model-uncertainty consequences. Contrast the results. This is easiest if the issue has only one primary dimension.[b]  
  *Example:* in estimating outcome of an adversary's internal debates, use one model that assumes outcome as restrained by compromise among factions and another that assumes the strident outcome of the winning faction. |
| Ensemble of Models    | Elicit diverse alternative models and consider all of them.  
  *Example:* in the study of climate-change, consider results of models from different universities, government laboratories, and private institutions.[c] |
| “Models” Inferred from Causal Statements | Require people to affirm important causal relationships among key variables comprising a system for which there is no formal model. Each set of statements acts as an alternative model, the differences among which have been formalized and are then subject to evaluation and discussion.  
  *Example:* in assessing the value of the U.S. supplying equipment to a regional security partner, require explicit statements about the benefits and drawbacks for such objectives as joint effectiveness and human and institutional capacity building. Ask for signs and magnitudes of effects for one unit of supply delivery and for many, and for the shape of the combined-effect curve (e.g., linear, decreasing returns, or sigmoid?) (O’Mahony et al., 2018). |
| Aggregate Parameters  | Forego modeling mechanisms and instead focus on aggregate consequences represented by parameters (a type of minimalist modeling).  
  *Examples:*  
  1) New technologies could mitigate effects of climate change but it is difficult to anticipate the changes or their effects. One could approach the issue at an aggregate level by varying a suitable parameter. Two such parameters are the decoupling rate between economic growth and growth of energy production and the decoupling rate between economic growth and growth of greenhouse gas emissions (Lempert et al., 2003).  
  2) Trade between Korea and China may be affected by the relative shares of GDP devoted to R&D but the exact mechanism and scale of change proxied by this value is just conjecture. Uncertainty over the magnitude of the effect may be represented with a multiplicative scalar between 0 and 1 (Seong & Popper, 2005). |

[a] An early example used analytic war games with competing "Red agents" to study Cold War deterrence (Davis, 1989). In 1990-1991, competing models of Saddam Hussein helped anticipate Saddam’s 1990 aggression against Kuwait and subsequent behavior (Davis & Arquilla, 1991; National Research Council, 2014). These efforts included effects of cognitive biases and misperceptions. Later, competing models of Kim Jong Il were used to study possible negotiation options. The study concluded that Kim was very unlikely (across plausible models) to truly give up his nuclear program (Arquilla & Davis, 1994). Most recently, competitive models of Kim Jong Un were used in contemplating the potential in a nuclear crisis...
(Davis, 2017), particularly the potential in crisis for escalation due to erroneous U.S. perceptions of Kim Jong Un and vice versa.

[b] The method was used for a model of public support for terrorism. Since the combined effects of relevant factors could be quite different depending on microscopic details, alternative sub models were employed to show the effect (Davis and O'Mahony, 2013). Another application experimented with heterogeneous fusion of information bearing on whether an individual should be judged as a potential terrorist threat. Results sometimes varied markedly depending on fusion method. A bounding-model approach proved useful, especially in illustrating analytically how false alarms can be given more credibility than they deserve (Davis, Perry, Hollywood, & Manheim, 2016).

[c] See Intergovernmental Panel on Climate Change (IPCC), 2014; Intergovernmental Panel on Climate Change (IPCC), et al.), 2010. Some see the result as demonstrating that such models are not yet good enough to be used to inform policy decisions (Corman, 2012), while others point out that the uncertainties do not "balance each other off" (see Chapter 2 of IPPC, 2014). Conclusions about the potentially catastrophic consequences of global warming are robust, while conclusions about the value of alternative interventions are less reliable.

Although overlapping to some degree, the approaches in Table 1 are conceptually rather distinct approaches to dealing with model uncertainty. To be sure, they may be mechanically implemented using a variant of parametric methods, as with giving the model a switch telling it which alternative model to use in a given run or constructing a composite model that takes on distinctly different characters depending on coefficients and parameter values (as with the Cobb-Douglas production function of economics).

A skeptical reader might politely ask at this point, "But how, with any of these methods, do we know that we have covered the range of possible model structures?" The answer, of course, is that we cannot be certain. Sometimes, mathematics or physical knowledge bounds the range of parameter values, but we have no such luxuries when addressing model uncertainty. We should aspire to "due diligence" in assuring that we consider model uncertainties of which we are aware, or should be aware after debate and the use of techniques such as in Table 1. That does not mean that we complicate analysis by tossing in every possibility someone can think of. To do would lead to paralysis. The practitioner analyst must employ a mix of art and science, and must subjectively omit some possibilities (invasions from Mars?). Nonetheless, our experience has been that pushing to address model uncertainty typically leads to better outcomes—i.e., assessments that better represent the best we can do with current knowledge.

4.2 Broadening the Concept of Analysis Campaign

A next element of a way ahead is broadening the concept of an analysis campaign, i.e., the concept of seeing an analysis project as something to be designed with appropriate attention paid to process, iteration, and other considerations. For example, such a campaign might have elements as in Table 2. Such a campaign applies to uncertainty analysis generally, but we have highlighted (with + marks) where it involves addressing model uncertainty in particular. We use the first column to indicate classic aspects of operations analysis (Greenberger, Crenson, & Crissey, 1976; Walker, 2000), the second column to note features that we now see
as coming into acceptance, and the third column is commentary to indicate additional features on which more emphasis is needed.

### Table 2 Increasingly Ambitious Treatments of Uncertainty

<table>
<thead>
<tr>
<th>Early</th>
<th>Advanced (primarily parametric uncertainty)</th>
<th>More Advanced, Routinely Addressing Model Uncertainty and Multiple Perspectives</th>
</tr>
</thead>
</table>
| Identify problem | Understand system and problem area | + Recognize parties, stakeholders, and views  
+ Understand broadened system, issues, questions, and tradeoffs  
+ Recognize emergent phenomena [a] |
| Build relatively narrow model | Start building evolving system model addressing phenomena, issues, and questions | + Include alternative model constructs  
+ Allow for alternative perspectives  
+ Represent emergent phenomena |
| Identify objectives and possible goals | Understand the multiple objectives (often in tension) and possible constraints of and on parties. Understand underlying values. | + Allow objectives to vary across models and perspectives  
+ Reflect in emergent phenomena |
| Choose evaluation criteria | Choose criteria, including "soft" criteria reflecting qualitative values. | + Allow criteria to vary across models and perspectives |
| Plan sensitivity analysis on a few primary variables | Plan broad parametric exploration across all significant variables | + Broaden exploration to cut across model structure and alternative perspectives |
| Select policy alternatives | Select or construct policy options, including creative options not offered initially | Include options that would allow for changes of system or processes |
| Analyze alternatives; compare by criteria above | Analyze with multi-criteria methods (e.g. scorecards). Do not "add up" scores prematurely, if ever. Retain visibility by major criteria. Evaluate across uncertainty space. | + Extend to address model uncertainties and alternative perspectives |
| Choose the best option (optimization) and implement it. | Identify options that are relatively robust to parametric uncertainties, i.e., options that are flexible, adaptive, and robust to shocks. Choose accordingly with expectation of adaptation later. | + Extend to reflect treatment of model uncertainty and alternative perspectives |
| Monitor results | Monitor and adapt, sometimes on the margin and sometimes with major shifts | + Extend to include adaptations changing primary model or perspective or changing balance thereof |

[a] This is crucial for "wicked problems," the norm in higher-level policy problem areas (Rosenhead & Mingers, 2002). How to represent emergence with generative models is a frontier challenge (Davis, O’Mahony, & Pfautz, 2019)

Another way to understand needs within an analysis campaign is to contemplate Figure 1 (Davis, 2014). This highlights the fact that (1) decisionmakers (plural) have views, perspectives, and disagreements that are not objectively resolved by analysis alone; (2) "other" considerations come into play with analysis being more able to inform some considerations than others; and (3) decisionmakers need not only to reach and report their decisions, but to also
explain them convincingly and provide meaningful guidance. Otherwise, decisions may not be accepted and implementation will fail.

**Figure 1 Analysis in a Decisionmaking Context**

Figure 2 suggests how an analysis campaign can affect results. It recognizes that analysis should have both deep and shallow components: deep to assure that critical phenomena are understood and shallow to permit good communication, explanation, and persuasion. Policymakers and those receiving decisions need to understand the reasoning. This requires simple conceptual models and compelling stories (alternative models/stories in some cases). A feature of the analysis campaign is that it anticipates being able to report the top-level story (or conflicting stories) but also recognizes the need to generate in-depth explanation as necessary. This means zooming into detail as necessary to support findings and recommendations.
Some other features of an analysis campaign should be breadth of scope and methods and of debate.

**Scope.** A good analysis campaign should draw on diverse sources for information and ideas. Some of this happens naturally in good research organizations. We have in mind models as diverse as statistical regressions, mathematical formulas, agent-based simulations, and live gaming. We also have in mind informing system concepts with insights from people as diverse as anthropologists, economists, and urban planners. This suggests that study may begin with a divergent phase in which a broad view is taken of "the system," the questions, the possible ways of modeling, and the crucial uncertainties and disagreements not to be lost—including different conceptions of the system itself.\(^7\)

**Debates.** A good study will typically uncover fundamentally different ways to understand the system of interest, the phenomena that occur within it, the cause-effect relationships, and the likely or possible result of interventions. This discord opens minds and allows a more objective analysis with the humbling influence of being sensitive to uncertainties and disagreements. In this environment, it will often be natural to contemplate analysis with different models, not just with different assumptions about model input.

As an example, suppose that the study is about whether to use economic "dynamic scoring" in evaluating federal legislation and, if so, how to do so. Dynamic scoring refers to using models to predict the impact of fiscal policy changes by forecasting the effects of economic agents' reactions to incentives created by macroeconomic policy. It is in contrast to static scoring that implicitly assumes no change in behavior. Upon beginning a good study on the subject, it would quickly become clear that major disagreements exist, even among reputable

---

\(^7\)Similar admonitions loom large in recommendations for improving the quality and usefulness of social-behavioral modeling (Davis, O'Mahony, & Pfautz, 2019).
economists. Those disagreements should be represented in qualitative models and then in computational versions. Some of the differences may be parametric, some will reflect model uncertainty, and some will be in a gray area.\(^8\)

Clearly, every effort should be made to resolve differences by exploiting empirical information. However, as experienced modelers know, that is not straightforward and people may reasonably draw opposite conclusions with the same data. Further, parties to a debate well understand that the initial assumptions may determine final outcomes and so will tenaciously resist early rejection of the views they hold. When this is the case, residual disagreements must be carried through the analysis.

**4.3 Improving the Communication of Insights**

All readers would probably agree that great efforts should be made to improve communication among modelers, analysts, and policymakers. It is notoriously difficult to do so and attempts are often humbling. Earlier, we mentioned the *Limits to Growth* controversy. The authors had labored to make their model and study understandable. We (the current authors) have done similarly in our own work, but have only sometimes succeeded. Nonetheless, much can be done and modern concepts and technology are making this increasingly feasible.

Some admonitions that we offer are:

- **Elevator Speech.** Be sure always to prepare the "elevator speech" on results and insights, but highlight model uncertainty usefully when appropriate.\(^9\)
- **Conceptual Model.** Increase emphasis on a distinct conceptual model, rather than the computerized model. Diagrams can play a major role in this, as they do in System Dynamics with other visual languages (e.g., *Analytica*), but in some cases the basic concepts are better described in tables (e.g., logic tables or game-theory tables) or even prose.
- **Multi-resolution Layering.** Convey complex models with a multiresolution layered approach: classic one-layer system-dynamic diagrams are too difficult to comprehend. Indeed, we find this to be good advice not only in the final presentation but in the analysis campaign itself as model uncertainty is explored through iterative analysis.
- **Qualitative Models.** Embrace qualitative models, especially for description and explanation. Examples include system diagrams, factor trees, influence diagrams, and social-network diagrams. A chronic problem in communication has been the analytic

---

\(^8\) As an example, opposing groups might agree that a Laffer Curve applies (i.e., that government revenue increases with tax rate, reaches some maximum, and then falls). They may disagree strongly about where the maximum occurs. Thus, one could say that they agree on model but not data. The effect, however, is so striking it is like using different models.

\(^9\) As an example, "Putting aside the analyses that just don't stand up, two respectable ways of viewing the problem remain. The first... The other version is... Decisions should hedge so that, as we get more information over time, we can accommodate to either."
community’s sometimes-exclusive and sometimes-misplaced emphasis on quantification.

- Tradeoffs, Not Answers. Emphasize tradeoff charts showing how outcomes vary with different combinations of causal-factor values and different model structures. This means that instead of showing results and waiting for the recipient of the work to say "Ok, but what if....?", the analysis preemptively shows results across the range of possibilities.

We illustrate this last approach and contrast in the next two figures. The top pane in Figure 3 shows a conventional comparison of two options for some standard case. The lower pane reveals much more: it shows under what conditions Option 1 or Option 2 is better. For simplicity, it uses only two dimensions (in this case, the parametric uncertainties of mission difficulty and timeliness required). It is apparent that Option 2, not Option 1, is better unless one is certain that only the “standard case” needs to be considered (no uncertainty). Other papers demonstrate methods for showing similar results with 6-10 independent variables (Davis, 2014) and methods for projecting multi-dimensional uncertainty analysis onto two dimensions (Lempert, Groves, Popper, & Bankes, 2006).

**Figure 3 Illustrating Region Plots**

![Figure 3 Illustrating Region Plots](image)

Figure 4 conveys a similar story for a more complex case. Imagine an analysis of alternative strategies such that, following intensive winnowing, modification and amalgamation, four fundamentally different strategies or plans remain, as does disagreement about the correct underlying model. Myriad cases have been generated for each plan with varying assumptions about which model applies, the results are aggregated and an expected outcome is calculated for each alternative depending on the odds of one model or the other prevailing. These outcomes are measured as the expected “regret” of having chosen each plan (regret is a
measure of how much better one might have done with a different decision) as a function of the likelihood that model M1 will prove more accurate. We see that Plan B performs best when the odds of Model M2 are greatest (left side) and Plan D has the least expected regret when Model M1 is most likely (right side). Plan A is dominated at every point by other plans, usually with large regret. Plan C shows interesting properties. If we are conducting an optimization analysis by assuming one model or the other, Plan C is unlikely to be of interest since at no point is it best (i.e., no matter which model is chosen, Plan C is not the best). But when we recognize model uncertainty, Plan C places a close second throughout and fails relatively gracefully with small expected regret. That is, it hedges well. Adopting Plan C could avoid a potentially costly blunder (the result of guessing wrong about which model to assume).

**Figure 4 Using a Measure of "Regret" to Assess an Option's Robustness to Assumptions**

Note: Because more regret is "bad," a plan is better if it lies lower rather than higher in this chart.

Earlier in the paper we mentioned that, historically, people (including decisionmakers) have had a hard time dealing with contingent predictions, one of the problems afflicting debate about *Limits to Growth*. The problems lie deep in human psychology, but we know from experience that analysts and the senior leaders to whom they report can learn to think in terms of tradeoff charts such as in Figure 3 and Figure 4. They are then internalizing the contingent nature of outcomes.

### 5 Existence Proofs

This paper stems from concern about analysis not routinely addressing model uncertainty. Nonetheless, enough examples exist in prior work to demonstrate that attempting to do so is often feasible and helpful. Without intending to exaggerate the definitiveness or persuasiveness of such past work, we mention the following.

1. Planning Military Capabilities. In recent decades the U.S. Department of Defense (DoD) has moved away from planning based on single scenarios to an approach that
seeks to assure capabilities adequate to deal with a broad range of possible conflicts and crises, and to do so while recognizing that in many cases it is difficult to estimate what the effectiveness of those capabilities will be because their effectiveness will depend on how the adversary behaves and fights and because some military strategies may prove either effective or counterproductive depending on contextual details, including the intended and perceived relationship with local factions. Although model uncertainties are particularly striking when dealing with counterterrorism and counterinsurgency, they loom large also in addressing more traditional force-sizing and force-posture analyses (Davis, Gompert, Johnson, & Long, 2008).

2. Studies of Counterterrorism and Counterinsurgency. The DoD confronted a painful reality in the late 2000's: its traditional models were unsuited for dealing with counterterrorism and counterinsurgency, subjects on which great disagreements exist about cause-effect relationships and the potential effectiveness of interventions. The DoD requested research "going back to basics," i.e., going back to study the underlying base in social science that should inform modeling and analysis. The resulting work generated mostly qualitative models with an emphasis on such models being able to deal with model uncertainty.

3. Industrial Transformation and Sustainability of Emission Control Policies. Although many aspects of climate change are largely agreed within the scientific community, others are not. Further, policy prescriptions remain highly controversial. Illuminating the issues is not possible by merely exercising standard economic models. To the contrary, understanding the issues requires thinking about possible transformational effects on commercial, energy, and transformation systems. One study did considerable alternative and speculative theorizing to improve policy debate.

4. International Investment in Clean-Energy Technology. The United Nations agreed in 2010 to the creation of the Green Climate Fund (GCF). A new organization was tasked with directing over $100B per year towards investments in clean energy technologies. Deep model uncertainties exist in attempting to understand these issues—

---

10 Early work informed the first Quadrennial Defense Review and the next administration's move to what DoD calls capabilities-based planning (Davis, Gompert, & Kugler, 1996; Davis, 2014).
11 Davis & Cragin, 2009 reviewed the social-science literature and constructed integrating conceptual models. Later work included qualitative validation of an integrative social-science model of public support for terrorism (Davis, Larson, Haldeman, Oguz, & Rana, 2012) and development of a subsequent uncertainty-sensitive computational model (Davis & O’Mahony, 2013; Davis & O’Mahony, 2017). Thus, the model was designed from the outset for exploring assumptions about model structure as well as parameter values.
12 The study used classic economics, game theory, simulation, and robust decisionmaking. Instead of focusing on prediction, the study discussed adaptive co-evolutionary possibilities for technology, political coalitions, industry, and government. It recognized the need for experiments and iterations of policy as the future unfolds (Isley, Lempert, Popper, & Vardavas, 2013).
i.e., about how the world works politically and economically, and about many aspects of climate change and mitigation methods. As shown in a recent study, however, much can be done to confront these and to find a robust set of policies (Molina-Perez, 2016).

5. Water Management Planning. Over the past decade, water resources managers and their partners have begun to use methods of decisionmaking under deep uncertainty (DMDU) to account for potential uncertain changes in hydrology and other drivers of supply and demand. Beginning with case studies in California (Groves & Lempert, 2007; Tingstad, Groves, & Lempert, 2014), these methods subsequently have been used to evaluate the vulnerability of complex water systems (Lempert & Groves, 2010; Groves, Fischbach, Bloom, Knopman, & Keefe, 2013). Researchers routinely use these or similar methods to evaluate water management issues internationally (Kalra et al., 2015; Ray et al., 2018).

6 Conclusions and Recommendations

Our basic recommendation is that policy analysts and those who commission their work should put a high priority on serious uncertainty analysis, to include addressing adequately the uncertainty of conceptual models that underlie the analysis. Doing so is a major challenge for analysts: it will mean providing analysis to help find strategies that are variously referred to as flexible, adaptive, and robust or—with the same meaning—to inform “robust decision-making.” That is, the strategy chosen should be those expected to do well across the range of uncertain assumptions—not just about inputs, but also about the models attempting to describe the way the world works and how alternative policies affect outcomes over time. In our view, the need to do so should become a basic ethic for analysts of choice under uncertainty, an ethic at least as important as that of revealing assumptions (Davis, 2014). For many years, analysts have understood that it is their professional responsibility to identify assumptions. The next step is taking responsibility for pointing out ways to hedge, whether or not asked.

Success in addressing the challenge will require basic cultural changes in both the analytic community and the community of policymakers commissioning and using analysis. These enterprises are intertwined. Policies are often promulgated with an accompanying fiction that issues are resolved by decision. The truth is different. A policy addressing a complex problem is more properly viewed as a policy experiment. It will cause new and observable behaviors, phenomena, and events. These should both provide more insight and suggest inevitable course corrections, sometimes major ones. Yet, not only do the policy makers ask their analysts questions embedding the fiction of policy finality, the analysts respond in ways that reinforce the fiction.

Table 3 characterizes the needed culture changes. The first column shows common current-day questions asked by policymakers and posed by those developing terms of reference for studies. The second column indicates the better types of question that would represent a shift in analytic culture.
### Table 3 Culture Changes Needed for Analysis Under Uncertainty

<table>
<thead>
<tr>
<th>Current Type of Question Asked by Policymakers</th>
<th>Better Types of Questions for Policy Under Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>What's your best prediction? What should we prepare for?</td>
<td>I know that reliable prediction is not in the cards. Thus, what are the things I can do now that will best position us to deal with whatever arises down the pike?</td>
</tr>
<tr>
<td>Which option is best?</td>
<td>How do I make a robustly good decision? What is necessary to have a strategy that is as flexible, adaptive, and robust as possible given budgetary and other considerations? What aspects of this will your analysis perform and what aspects will be left for me to worry about on my own?</td>
</tr>
<tr>
<td>Whose options are you going to look at?</td>
<td>Where will your options come from? Are you talking to everyone? Are you going to construct options to cover what needs to be covered when the bureaucracy doesn't do so by itself? Will there be new ideas?</td>
</tr>
<tr>
<td>What data are you using?</td>
<td>What is your campaign plan for analysis? Does it assure that you will be addressing all the fundamental uncertainties and disagreements, as well as the uncertainties of so-called data (a mixture of empirical data and speculative assumptions)?</td>
</tr>
<tr>
<td>Do you have a steering group with reps from the relevant offices?</td>
<td>Do you plan to have a “red team” apart from your analysis team who will ask the hard questions? Do you plan to have an Advisory Group to assure that you're covering all the bases, a Group of not just the “official” stakeholders, but of everyone who should be heard? Will implementation be part of the campaign of analysis or will that be an afterthought?</td>
</tr>
<tr>
<td>Are you using a validated model and data?</td>
<td>How is your campaign plan for analysis going to assure that your modeling and analysis are valid for the purposes intended, whether to &quot;describe and explain,&quot; to &quot;explore to find insights and the ingredients of robust decisions,&quot; or prediction? What model(s) are you going to consider? Are there competing models that should be considered?</td>
</tr>
<tr>
<td>Can I ask What if? questions about results under different scenarios?</td>
<td>Are you going to give me intelligible big-picture estimates as a function of major variables? Tradeoff plots? Measures of robustness (e.g., regret)? Will I have to just listen or can I interact? Will there be real-time responses to questions or requests for another six months/</td>
</tr>
<tr>
<td>How long do you need for your brief-out? 30 minutes? More?</td>
<td>How am I going to interact along the way? Initial plan, interim discussion, final discussion? Are you going to have results in layers of detail so that I (or my staff) can get into details where necessary? Will the study be documented intelligibly?</td>
</tr>
<tr>
<td>Oh, how do I monitor to assure that my decision is implemented?</td>
<td>Oh, since we’ll learn from experience (and perhaps recognize error), how should we plan to monitor and adapt?</td>
</tr>
</tbody>
</table>

It is not the role of analysis and analysts to create deus ex machina in the form of analyses that will resolve our policy choices. That is rightly the job of elected officials and their appointees. There are always ancillary questions of politics and policy, details and interests, that must as a matter of course also enter the policy deliberation process. But what analysts can do is illuminate the policy trade-off space: which short-term actions (and preparation for later adaptations) will best achieve our long-term policy objectives despite broad uncertainties now and the likelihood of unexpected developments as the future unfolds? How does the an-
swer depend on matters on which there is disagreement? If analysts can do this, they may help bring about the needed shift in cultures.

**Acknowledgments**

This paper was possible due to support by the RAND Center for Decision Making under Uncertainty, within the Pardee RAND Graduate School. We thank Andrew Parker and David Groves for comments on the first draft.

**Bibliography**


technical_reports/TR1220.html


